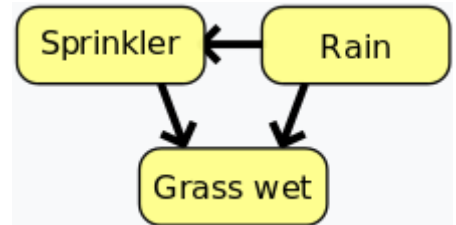


貝氏網路

維基百科，自由的百科全書

貝氏網路（Bayesian network），又稱**信念網絡**（belief network）或是**有向無環圖模型**（directed acyclic graphical model），是一種機率圖型模型，藉由**有向無環圖**（directed acyclic graphs, or DAGs）中得知一組隨機變數 $\{X_1, X_2, \dots, X_n\}$ 及其 n 組**條件機率分配**（conditional probability distributions, or CPDs）的性質。舉例而言，貝氏網路可用來表示疾病和其相關症狀間的機率關係；倘若已知某種症狀下，貝氏網路就可用來計算各種可能罹患疾病之發生機率。



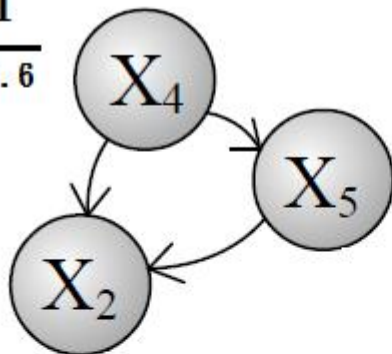
一個簡單的貝氏網路。雨水影響灑水器是否有動作，且雨水及灑水器二者均可影響草是否濕潤。

一般而言，貝氏網路的有向無環圖中的節點表示隨機變數，它們可以是可觀察到的變量，抑或是潛在變量、未知參數等。連接兩個節點的箭頭代表此兩個隨機變數是具有因果關係或是非條件獨立的；而兩個節點間若沒有箭頭相互連接一起的情況就稱其隨機變數彼此間為條件獨立。若兩個節點間以一個單箭頭連接在一起，表示其中一個節點是「**因**（parents）」，另一個是「**果**（descendants or children）」，兩節點就會產生一個條件機率值。比方說，我們以 X_i 表示第 i 個節點，而 X_i 的「因」以 P_i 表示， X_i 的「果」以 C_i 表示；圖一就是一種典型的貝氏網路結構圖，依照先前的定義，我們就可以輕易的從圖一可以得知：

$$P_2 = \{X_4, X_5\}, C_2 = \emptyset, P_4 = \emptyset, C_4 = \{X_2, X_5\}, P_5 = \{X_4\}, \text{ 以及 } C_5 = \{X_2\}$$

大部分的情況下，貝氏網路適用在節點的性質是屬於離散型的情況下，且依照 $P(X_i|P_i)$ 此條件機率寫出**條件機率表**（conditional probability table, or CPT），此條件機率表的每一列（row）列出所有可能發生的 P_i ，每一行（column）列出所有可能發生的 X_i ，且任一列的機率總和必為1。寫出條件機率表後就很容易將事情給條理化，且輕易地得知此貝氏網路結構圖中各節點間之因果關係；但是條件機率表也有其缺點：若是節點 X_i 是由很多的「因」所造成的「果」，如此條件機率表就會變得在計算上既複雜又使用不便。下圖為圖一貝氏網路中某部分結構圖之條件機率表。

X_4	0	1
prob.	0.4	0.6



X_4

	X_5	
	0	1
0	0.32	0.68
1	0.54	0.46

		X_2	
		0	1
X_5	X_4		
0	0	1.0	0.0
0	1	0.2	0.8
1	0	0.91	0.09
1	1	0.02	0.98

圖一：部分結構圖之條件機率表

目录

[數學定義](#)

[馬可夫毯 \(Markov blanket\)](#)

[舉例說明](#)

[求解方法](#)

[精確推論](#)

[隨機推論 \(蒙地卡羅方法\)](#)

1. [結構已知，觀測值完整：](#)

2. [結構已知，觀測值不完整 \(有遺漏資料\)：](#)

[補充例子 \(列舉推理法\)](#)

[應用](#)

[參考文獻](#)

[外部連結](#)

數學定義

令 $G = (I, E)$ 表示一個有向無環圖 (DAG)，其中 I 代表圖中所有的節點的集合，而 E 代表有向連接線段的集合，且令 $X = (X_i)_{i \in I}$ 為其有向無環圖中的某一節點 i 所代表之隨機變數，若節點 X 的聯合機率分配可以表示成：

$$p(x) = \prod_{i \in I} p(x_i | x_{pa(i)})$$

則稱 X 為相對於一有向無環圖 G 的貝氏網路，其中 $pa(i)$ 表示節點 i 之「因」。

對任意的隨機變數，其聯合分配可由各自的局部條件機率分配相乘而得出：

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i | X_{i+1} = x_{i+1}, \dots, X_n = x_n)$$

依照上式，我們可以將一貝氏網路的聯合機率分配寫成：

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i | X_j = x_j) \text{ (對每個相對於 } X_i \text{ 的「因」變數 } X_j \text{ 而言)}$$

上面兩個表示式之差別在於條件機率的部分，在貝氏網路中，若已知其「因」變數下，某些節點會與其「因」變數條件獨立，只有與「因」變數有關的節點才会有條件機率的存在。

如果聯合分配的相依數目很稀少時，使用貝氏函數的方法可以節省相當大的記憶體容量。舉例而言，若想將10個變數其值皆為0或1儲存成一條件機率表型式，一個直觀的想法可知我們總共必須要計算 $2^{10} = 1024$ 個值；但若這10個變數中無任何變數之相關「因」變數是超過三個以上的話，則貝氏網路的條件機率表最多只需計算 $10 * 2^3 = 80$ 個值即可。另一個貝氏網路優點在於：對人類而言，它更能輕易地得知各變數間是否條件獨立或相依與其**局部分配** (local distribution) 的型態來求得所有隨機變數之聯合分配。

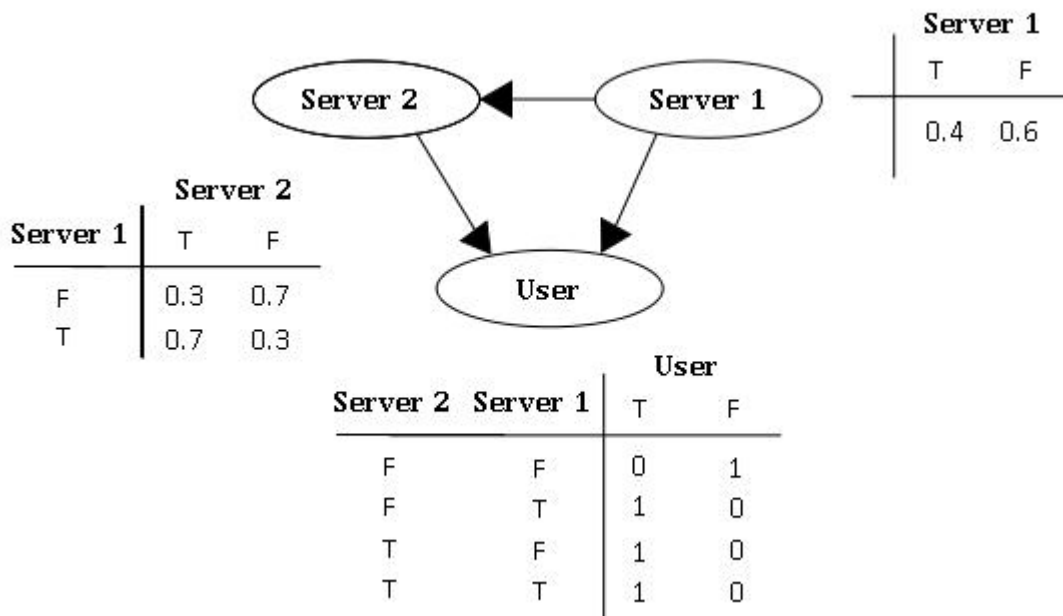
馬可夫毯 (Markov blanket)

定義一個節點之馬可夫毯為此節點的因節點、果節點與果節點的因節點所成之集合。一旦給定其馬可夫毯的值後，若網路內之任一節點 X 皆會與其他的節點條件獨立的話，就稱 X 為相對於一有向無環圖 G 的貝氏網路。

舉例說明

假設有兩個伺服器 (S_1, S_2)，會傳送封包到使用者端 (以 U 表示之)，但是第二個伺服器的封包傳送成功率會與第一個伺服器傳送成功與否有關，因此此貝氏網路的結構圖可以表示成如圖二的型式。就每個封包傳送而言，只有兩種可能值： T (成功) 或 F (失敗)。則此貝氏網路之聯合機率分配可以表示成：

$$P(U, S_1, S_2) = P(U | S_1, S_2) \times P(S_2 | S_1) \times P(S_1)$$



圖二：簡單的貝氏網路例子

此模型亦可回答如：「假設已知使用者端成功接受到封包，求第一伺服器成功發送封包的機率？」諸如此類的問題，而此類型問題皆可用條件機率的方法來算出其所求之發生機率：

$$\begin{aligned}
 P(S_1 = T \mid U = T) &= \frac{P(S_1 = T, U = T)}{P(U = T)} \\
 &= \frac{\sum_{S_2 \in \{T, F\}} P(S_1 = T, S_2, U = T)}{\sum_{S_1, S_2 \in \{T, F\}} P(U = T, S_1, S_2)} \\
 &= \frac{(0.4 \times 0.3 \times 1)_{U=T, S_1=T, S_2=F} + (0.4 \times 0.7 \times 1)_{U=T, S_1=T, S_2=T}}{(0.4 \times 0.3 \times 1)_{TTT} + (0.4 \times 0.7 \times 1)_{TTF} + (0.6 \times 0.3 \times 1)_{FTT} + 0_{TFF}} \\
 &\approx 68.96\%。
 \end{aligned}$$

求解方法

以上例子是一個很簡單的貝氏網路模型，但是如果當模型很複雜時，這時使用列舉式的方法來求解機率就會變得非常複雜且難以計算，因此必須使用其他的替代方法。一般來說，貝氏機率有以下幾種求法：

精確推論

- 列舉推理法（如上述例子）
- 變數消元演算法（variable elimination）

隨機推論（蒙地卡羅方法）

- 直接取樣演算法
- 拒絕取樣演算法
- 概似加權演算法

■ 馬可夫鏈蒙地卡羅演算法 (Markov chain Monte Carlo algorithm)

在此，以馬可夫鏈蒙地卡羅演算法為例，又馬可夫鏈蒙地卡羅演算法的類型很多，故在這裡只說明其中一種吉布斯採樣的操作步驟：首先將已給定數值的變數固定，然後將未給定數值的其他變數隨意給定一個初始值，接著進入以下迭代步驟：

- (1) 隨意挑選其中一個未給定數值的變數
- (2) 從條件分配 $P(X_i | \text{Markovblanket}(X_i))$ 抽樣出新的 X_i 的值，接著重新計算

$$P(X_i | \text{Markovblanket}(X_i)) = \alpha P(X_i | \text{parents}(X_i)) \times \prod_{Y_i \in \text{children}(X_i)} \text{parent}(Y_i)$$

當迭代結束後，刪除前面若干筆尚未穩定的數值，就可以求出的近似條件機率分配。馬可夫鏈蒙地卡羅演算法的優點是在計算很大的網路時效率很好，但缺點是所抽取出的樣本並不具獨立性。

當貝氏網路上的結構跟參數皆已知時，我們可以透過以上方法來求得特定情況的機率，不過，如果當網路的結構或參數未知時，我們必須藉由所觀測到的資料去推估網路的結構或參數，一般而言，推估網路的結構會比推估節點上的參數來的困難。依照對貝氏網路結構的了解和觀測值的完整與否，我們可以分成下列四種情形：

結構	觀測值	方法
已知	完整	最大概似估計法 (MLE)
已知	部份	EM演算法 Greedy Hill-climbing method
未知	完整	搜尋整個模型空間
未知	部份	結構演算法 EM演算法 Bound contraction

以下就結構已知的部分，作進一步的說明。

1. 結構已知，觀測值完整：

此時我們可以用最大概似估計法 (MLE) 來求得參數。其對數概似函數為

$$L = \frac{1}{N} \sum_{i=1}^n \sum_{j=1}^s \log(P(X_i | \text{pa}(X_i), D_i))$$

其中 $\text{pa}(X_i)$ 代表 X_i 的因變數， D_i 代表第 i 個觀測值， N 代表觀測值資料的總數。

以圖二當例子，我們可以求出節點 U 的最大概似估計式為

$$P(U = u | S_1 = s_1, S_2 = s_2) = \frac{n(U = u, S_1 = s_1, S_2 = s_2)}{n(S_1 = s_1, S_2 = s_2)}$$

由上式就可以藉由觀測值來估計出節點 U 的條件分配。如果當模型很複雜時，這時可能就要利用數值分析或其它最佳化技巧來求出參數。

2. 結構已知，觀測值不完整 (有遺漏資料)：

如果有些節點觀測不到的話，可以使用EM演算法 (*Expectation-Maximization algorithm*) 來決定出參數的區域最佳概似估計式。而EM演算法的主要精神在於如果所有節點的值都已知下，在M階段就會很簡單，如同最大概似估計法。而EM演算法的步驟如下：

(1) 首先給定欲估計的參數一個起始值，然後利用此起始值和其他的觀測值，求出其他未觀測到節點的條件期望值，接著將所估計出的值視為觀測值，將此完整的觀測值帶入此模型的最大概似估計式中，如下所示 (以圖二為例)：

$$P(U = u | S_1 = s_1, S_2 = s_2) = \frac{EN(U = u, S_1 = s_1, S_2 = s_2)}{EN(S_1 = s_1, S_2 = s_2)}$$

其中 $EN(x)$ 代表在目前的估計參數下，事件 x 的條件機率期望值為

$$EN(x) = E \sum_k I(x|D(k)) = \sum_k P(x|D(k))$$

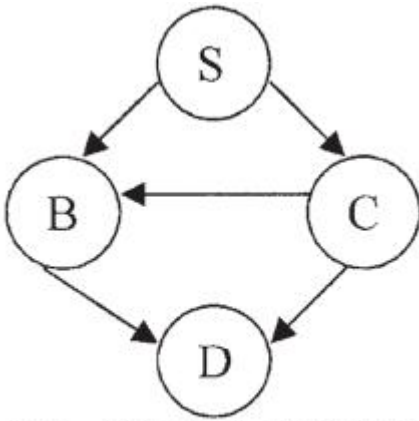
(2) 最大化此最大概似估計式，求出此參數之最有可能值，如此重複步驟 (1) 與 (2)，直到參數收斂為止，即可得到最佳的參數估計值。

補充例子 (列舉推理法)

讓我們考慮一個應用在醫藥上的機率推論例子，在此病人會被診斷出是否有呼吸困難的症狀。表一代表一個我們所觀測到的資料集合，包含10筆觀測值， S 代表的是吸菸與否 (Smoker)， C 代表是否為罹癌者 (Cancer)， B 代表是否罹患支氣管炎 (bronchitis)， D 代表是否有呼吸困難及咳嗽 (dyspnea and asthma) 的症狀。『1』和『0』分別代表『是』和『否』。此醫藥網路結構顯示於圖三。

表一

No. of observation	Answers to learning data			
	S	C	B	D
1	1	1	1	1
2	1	0	1	0
3	1	0	0	0
4	0	0	1	1
5	0	0	0	0
6	1	0	0	0
7	1	1	1	1
8	0	0	0	0
9	0	0	0	1
10	0	0	0	0



圖三：醫藥的貝氏網路結構

表二代表的是整個網路的經驗聯合機率分配，是由所收集到的資料所建構而成，利用此表可建構出節點的聯合機率分配。見圖四。此貝氏公式 $P(A|B) = \frac{P(A, B)}{P(B)}$ 可利用節點的邊際機率和聯合機率去計算節點的條件機率，待會會應用在建立條件機率表格 (Conditional probability Table; CPT) 上。見圖五。貝氏網路的聯合機率可由下列式子計算：

$$P(S, C, B, D) = P(S) \times P(C|S) \times P(B|S, C) \times P(D|B, C)$$

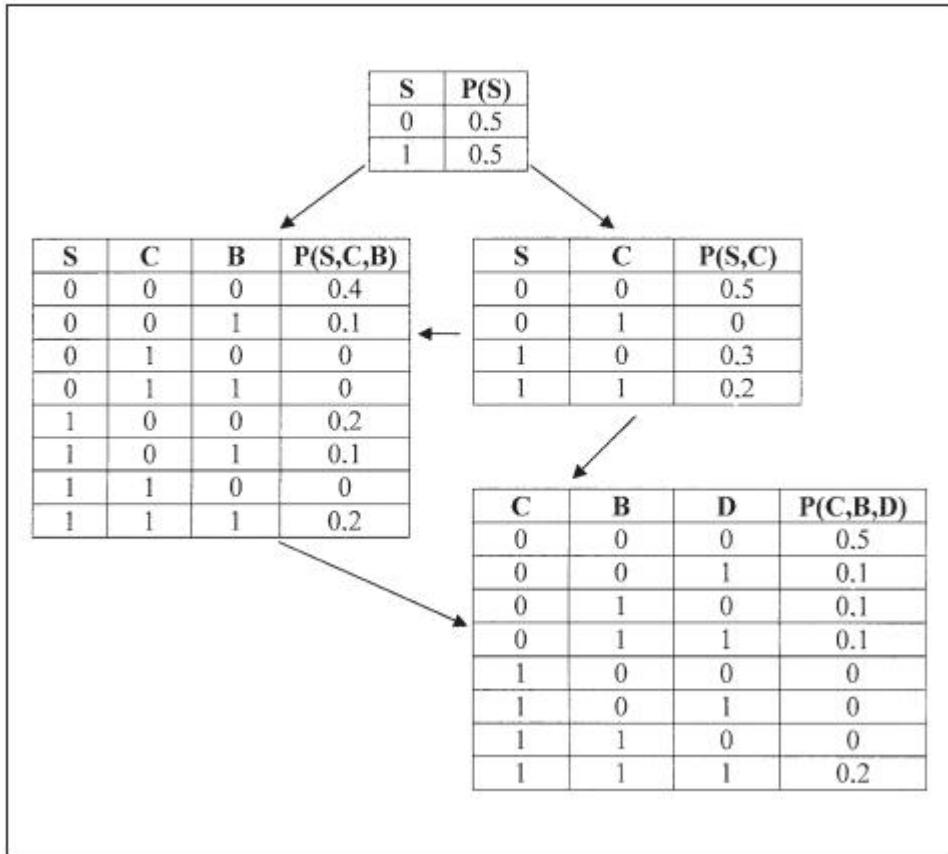
其值見表三。

使用整個網路經驗聯合機率分配所計算出來的值會與使用CPT所計算出來的值不同，其差異可由表二和表三得知。其中差異不只是值的不同，也出現了新事件的機率（原本所沒觀察到的事件）。

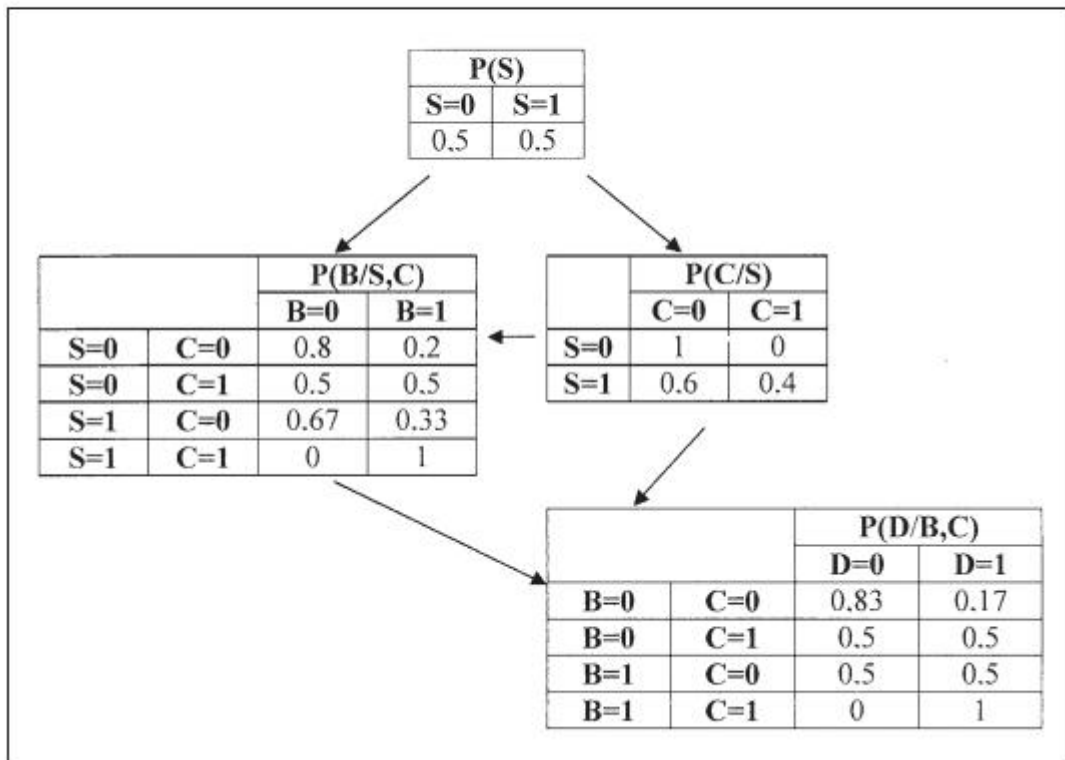
表二

No. of data set	Answers to learning data				Network probability
	<i>S</i>	<i>C</i>	<i>B</i>	<i>D</i>	
1	0	0	0	0	0.3
2	0	0	0	1	0.1
3	0	0	1	1	0.1
4	1	0	0	0	0.2
5	1	0	1	0	0.1
6	1	1	1	1	0.2

圖四



圖五



表三

No. of data set	Answers to learning data				Network probability
	<i>S</i>	<i>C</i>	<i>B</i>	<i>D</i>	
1	0	0	0	0	0.332
2	0	0	0	1	0.068
3	0	0	1	0	0.05
4	0	0	1	1	0.05
5	0	1	0	0	0
6	0	1	0	1	0
7	0	1	1	0	0
8	0	1	1	1	0
9	1	0	0	0	0.164
10	1	0	0	1	0.034
11	1	0	1	0	0.051
12	1	0	1	1	0.051
13	1	1	0	0	0
14	1	1	0	1	0
15	1	1	1	0	0
16	1	1	1	1	0.2

建立在觀測資料上的機率推論演算法：

1.資料集合 $D = \{d_1, \dots, d_n\}$ ，其中 $d_i = x_i^{(1)}, \dots, x_i^{(N)}$ （下標代表第幾個觀測值，上標代表第幾個變數），且一共有n個觀測值。每一個觀測值包含 $N(N \geq 2)$ 個變數 $X_i^{(1)}, \dots, X_i^{(N)}$ ，第j個變數有 $A^{(j)} = 0, 1, \dots, \alpha^{(j)} - 1 (\alpha^{(j)} \geq 2)$ 狀態。

2.此貝氏網路的結構G代表N個前代（predecessors）節點集合 $\{\Pi^{(1)}, \dots, \Pi^{(N)}\}$ ，也就是對第j個節點， $\Pi^{(j)}$ 為其親代節點的集合, $j=1, 2, \dots, N$

3.範例點（instantiated node） $\{X^{P_1} = x^{P_1}, \dots, X^{P_v} = x^{P_v}\}$ 為節點在已知狀態，即在此狀態的機率為1。如果範例點為空集合，將使用古典機率推論

使用表一的觀測值和圖一的貝氏網路結構，並且已知範例點（instantiated node）為 $\{S = 0, C = 0\}$ ，也就是病人為非吸菸者和罹癌者： $P(S = 0) = 1, P(C = 0) = 1$

問題：

- 1.病人患有支氣管炎的機率 $P(B) = ?$
- 2.病人會有呼吸困難的機率 $P(D) = ?$

解答：

1. $P(B = 0 | S = 0, C = 0) = 0.8$

$P(B = 1 | S = 0, C = 0) = 0.2$

故為0.2

$$2. P(D = 0|S = 0, C = 0)$$

$$\begin{aligned} &= P(D = 0, B = 0|S = 0, C = 0) + P(D = 0, B = 1|S = 0, C = 0) \\ &= P(D = 0|B = 0, S = 0, C = 0) \times P(B = 0|S = 0, C = 0) \\ &\quad + P(D = 0|B = 1, S = 0, C = 0) \times P(B = 1|S = 0, C = 0) \end{aligned}$$

$$P(D = 1|S = 0, C = 0)$$

$$\begin{aligned} &= P(D = 1, B = 1|S = 0, C = 0) + P(D = 1, B = 0|S = 0, C = 0) \\ &= P(D = 1|B = 0, S = 0, C = 0) \times P(B = 0|S = 0, C = 0) \\ &\quad + P(D = 1|B = 1, S = 0, C = 0) \times P(B = 1|S = 0, C = 0) \end{aligned}$$

應用

貝氏網路目前應用在類比計算生物學與生物資訊學基因調控網路、蛋白質結構、基因表達分析、醫學、文档分类、資訊檢索、決策支援系統、工程學、資料結合、圖像處理等。

參考文獻

- A. N. Terent' yev and P. I. Bidyuk. METHOD OF PROBABILISTIC INFERENCE FROM LEARNING DATA IN BAYESIAN NETWORKS. Cybernetics and Systems Analysis, 391-396, Vol. 43, No. 3, 2007
- Emilia Mendes and Nile Mosley. Bayesian Network Models for Web Effort Prediction: A Comparative Study. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, 723-737, VOL. 34, NO. 6, NOVEMBER/DECEMBER 2008
- Yu. V. Kapitonova, N. M. Mishchenko, O. D. Felizhanko, and N. N. Shchegoleva. USING BAYESIAN NETWORKS FOR MONITORING COMPUTER USERS. Cybernetics and Systems Analysis, 789-799, Vol. 40, No. 6, 2004
- Ben-Gal I., Bayesian Networks, in Ruggeri F., Faltin F. & Kenett R. ,Encyclopedia of Statistics in Quality & Reliability, Wiley & Sons (2007).
- Radu Stefan Niculescu, Tom M. Mitchell and R. Bharat Rao, Journal of Machine Learning Research 7 (2006) 1357–1383
- N. Terent' yev and P. I. Bidyuk. METHOD OF PROBABILISTIC INFERENCE FROM LEARNING DATA IN BAYESIAN NETWORKS. Cybernetics and Systems Analysis, 391-396, Vol. 43, No. 3, 2007
- P. I. Bidyuk, A. N. Terent'ev, and A. S. Gasanov. CONSTRUCTION AND METHODS OF LEARNING OF BAYESIAN NETWORKS. Cybernetics and Systems Analysis, 587-598, Vol. 41, No. 4, 2005

外部連結

- A tutorial on learning with Bayesian Networks (https://web.archive.org/web/20060719171558/http://research.microsoft.com/research/pubs/view.aspx?msr_tr_id=MSR-TR-95-06)
- An Introduction to Bayesian Networks and their Contemporary Applications (<http://www.niedermayer.ca/papers/bayesian/bayes.html>) (页面存档备份 (<https://web.archive.org/web/20170521075321/http://www.niedermayer.ca/papers/bayesian/bayes.html>), 存于互联网档案馆)
- On-line Tutorial on Bayesian nets and probability (<http://www.dcs.qmw.ac.uk/%7Enorman/BBNs/BBNs.htm>) (页面存档备份 (<https://web.archive.org/web/20090504220301/http://www.dcs>

s.gmw.ac.uk/%7Enorman/BBNs/BBNs.htm), 存于互联网档案馆)

- [Web-App to create Bayesian nets and run it with a Monte Carlo method \(https://web.archive.org/web/20170601002137/http://princesofserendib.com/\)](https://web.archive.org/web/20170601002137/http://princesofserendib.com/)
- [Continuous Time Bayesian Networks \(http://robotics.stanford.edu/~nodelman/papers/ctbn.pdf\)](http://robotics.stanford.edu/~nodelman/papers/ctbn.pdf) (页面存档备份 (https://web.archive.org/web/20181009114501/http://robotics.stanford.edu/~nodelman/papers/ctbn.pdf), 存于互联网档案馆)
- [Bayesian Networks: Explanation and Analogy \(https://web.archive.org/web/20131225224701/http://wiki.synclous.com/index.php/dANN:Bayesian_Network\)](https://web.archive.org/web/20131225224701/http://wiki.synclous.com/index.php/dANN:Bayesian_Network)
- [A live tutorial on learning Bayesian networks \(http://videlectures.net/kdd07_neapolitan_lbn/\)](http://videlectures.net/kdd07_neapolitan_lbn/) (页面存档备份 (https://web.archive.org/web/20210415003910/http://videlectures.net/kdd07_neapolitan_lbn/), 存于互联网档案馆)
- [A hierarchical Bayes Model for handling sample heterogeneity in classification problems \(http://www.biomedcentral.com/1471-2105/7/514/abstract\)](http://www.biomedcentral.com/1471-2105/7/514/abstract) (页面存档备份 (https://web.archive.org/web/20150109031607/http://www.biomedcentral.com/1471-2105/7/514/abstract), 存于互联网档案馆), provides a classification model taking into consideration the uncertainty associated with measuring replicate samples.
- [Hierarchical Naive Bayes Model for handling sample uncertainty \(http://www.labmedinfo.org/download/lmi339.pdf\)](http://www.labmedinfo.org/download/lmi339.pdf) (页面存档备份 (https://web.archive.org/web/20070928081740/http://www.labmedinfo.org/download/lmi339.pdf), 存于互联网档案馆), shows how to perform classification and learning with continuous and discrete variables with replicated measurements.

取自“<https://zh.wikipedia.org/w/index.php?title=貝氏網路&oldid=70614629>”

本页面最后修订于2022年3月14日 (星期一) 02:05。

本站的全部文字在知识共享 署名-相同方式共享 3.0协议之条款下提供, 附加条款亦可能应用。 (请参阅使用条款)

Wikipedia®和维基百科标志是维基媒体基金会的注册商标; 维基™是维基媒体基金会的商标。

维基媒体基金会是按美国国内稅收法501(c)(3)登记的非营利慈善机构。